PP

HPFE

Access DB# 27652

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: Nirmal S. Basi Examiner # : ________ Date: 10/24/00
Art Unit: 1646 Phone Number 308-9435 Serial Number: 09/125635
Mail Box and Bldg/Room Location: CM1 10E17 Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

****

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: AIB1 A novel Steroid Receptor

Inventors (please provide full names): Meltzer et al

Earliest Priority Filing Date: 6/16/97

***For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.**

Please search SEQ ID Nos: 1,2,3,4,5,6,7,8.

Commercial + interference search

PRO: Nuc

FOR OFFICIAL USE ONLY

****

STAFF USE ONLY		Type of Search		Vendors and cost where applicable	
Searcher:	B Smith	NA Sequence (#)		STN	
Searcher Phone #:	308-4497	AA Sequence (#)		Dialog	
Searcher Location:	CM1-1E17	Structure (#)		Questel/Orbit	
Date Searcher Picked Up:	10/24/00	Bibliographic		Dr.Link	
Date Completed:	11/2/00	Litigation		Lexis/Nexis	
Searcher Prep & Review Time:		Fulltext		Sequence Systems	AbSS03
Clerical Prep Time:	5	Patent Family		WWW/Internet	
Online Time:		Other		Other (specify)	

PTO-1590 (1-2000)

```
Qy   2936 tcagactccttcctcaggagactggggcttaccaaactcaaaggccggcagaatggaacc 2995
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2821 TCAGACTCCTTCCTCAGGAGACTGGGGCTTACCAAACTCAAAGGCCGGCAGAATGGAACC 2880

Qy   2996 tatgaattcaaactccatgggaagaccaggaggagattataatacttctttacccagacc 3055
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2881 TATGAATTCAAACTCCATGGGAAGACCAGGAGGAGATTATAATACTTCTTTACCCAGACC 2940

Qy   3056 tgcactgggtggctctattcccacattgcctcttcggtctaatagcataccaggtgcgag 3115
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2941 TGCACTGGGTGGCTCTATTCCCACATTGCCTCTTCGGTCTAATAGCATACCAGGTGCGAG 3000

Qy   3116 accagtattgcaacagcagcagcagatgcttcaaatgaggcctggtgaaatccccatggg 3175
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3001 ACCAGTATTGCAACAGCAGCAGCAGATGCTTCAAATGAGGCCTGGTGAAATCCCCATGGG 3060

Qy   3176 aatgggggctaatccctatggccaagcagcagcatctaaccaactgggttcctggcccga 3235
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3061 AATGGGGGCTAATCCCTATGGCCAAGCAGCAGCATCTAACCAACTGGGTTCCTGGCCCGA 3120

Qy   3236 tggcatgttgtccatggaacaagtttctcatggcactcaaaataggcctcttcttaggaa 3295
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3121 TGGCATGTTGTCCATGGAACAAGTTTCTCATGGCACTCAAAATAGGCCTCTTCTTAGGAA 3180

Qy   3296 ttccctggatgatcttgttgggccaccttccaacctggaaggccagagtgacgaaagagc 3355
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3181 TTCCCTGGATGATCTTGTTGGGCCACCTTCCAACCTGGAAGGCCAGAGTGACGAAAGAGC 3240

Qy   3356 attattggaccagctgcacactcttctcagcaacacagatgccacaggcctggaagaaat 3415
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3241 ATTATTGGACCAGCTGCACACTCTTCTCAGCAACACAGATGCGACAGGCCTGGAAGAAAT 3300

Qy   3416 tgacagagctttgggcattcctgaacttgtcaatcagggacaggcattagagcccaaaca 3475
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3301 TGACAGAGCTTTGGGCATTCCTGAACTTGTCAATCAGGGACAGGCATTAGAGCCCAAACA 3360

Qy   3476 ggatgctttccaaggccaagaagcagcagtaatgatggatcagaaggcaggattatatgg 3535
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3361 GGATGCTTTCCAAGGCCAAGAAGCAGCAGTAATGATGGATCAGAAGGCAGGATTATATGG 3420

Qy   3536 acagacatacccagcacaggggcctccaatgcaaggaggctttcatcttcagggacaatc 3595
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3421 ACAGACATACCCAGCACAGGGGCCTCCAATGCAAGGAGGCTTTCATCTTCAGGGACAATC 3480

Qy   3596 accatcttttaactctatgatgaatcagatgaaccagcaaggcaattttcctctccaagg 3655
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3481 ACCATCTTTTAACTCTATGATGAATCAGATGAACCAGCAAGGCAATTTTCCTCTCCAAGG 3540

Qy   3656 aatgcacccacgagccaacatcatgagaccccggacaaacacccccaagcaacttagaat 3715
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3541 AATGCACCCACGAGCCAACATCATGAGACCCCGGACAAACACCCCCAAGCAACTTAGAAT 3600

Qy   3716 gcagcttcagcagaggctgcagggccagcagtttttgaatcagagccgacaggcacttga 3775
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3601 GCAGCTTCAGCAGAGGCTGCAGGGCCAGCAGTTTTTGAATCAGAGCCGACAGGCACTTGA 3660

Qy   3776 attgaaaatggaaaaccctactgctggtggtgctgcggtgatgaggcctatgatgcagcc 3835
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3661 ATTGAAAATGGAAAACCCTACTGCTGGTGGTGCTGCGGTGATGAGGCCTATGATGCAGCC 3720

Qy   3836 ccagcagggttttcttaatgctcaaatggtcgcccaacgcagcagagagctgctaagtca 3895
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3721 CCAGCAGGGTTTTCTTAATGCTCAAATGGTCGCCCAACGCAGCAGAGAGCTGCTAAGTCA 3780

Qy   3896 tcacttccgacaacagagggtggctatgatgatgcagcagcagcagcagcagcaacagca 3955
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3781 TCACTTCCGACAACAGAGGGTGGCTATGATGATGCAGCAGCAGCAGCAGCAGCAACAGCA 3840

Qy   3956 gcagcagcagcagcagcagcagcaacagcaacagcaacagcaacagcagcaacagcagca 4015
          |||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db   3841 GCAGCAGCAGCAG---------CAGCAACAGCAACAGCAACAGCAGCAACAGCAGCA 3891
```

```
Db   3892 AACCCAGGCCTTCAGCCCACCTCCTAATGTGACTGCTTCCCCCAGCATGGATGGGCTTTT 3951
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   4076 ggcaggacccacaatgccacaagctcctccgcaacagtttccatatcaaccaaattatgg 4135
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3952 GGCAGGACCCACAATGCCACAAGCTCCTCCGCAACAGTTTCCATATCAACCAAATTATGG 4011

Qy   4136 aatgggacaacaaccagatccagcctttggtcgagtgtctagtcctcccaatgcaatgat 4195
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4012 AATGGGACAACAACCAGATCCAGCCTTTGGTCGAGTGTCTAGTCCTCCCAATGCAATGAT 4071

Qy   4196 gtcgtcaagaatgggtccctcccagaatcccatgatgcaacacccgcaggctgcatccat 4255
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4072 GTCGTCAAGAATGGGTCCCTCCCAGAATCCCATGATGCAACACCCGCAGGCTGCATCCAT 4131

Qy   4256 ctatcagtcctcagaaatgaagggctggccatcaggaaatttggccaggaacagctccttt 4315
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4132 CTATCAGTCCTCAGAAATGAAGGGCTGGCCATCAGGAAATTTGGCCAGGAACAGCTCCTT 4191

Qy   4316 ttcccagcagcagtttgcccaccaggggaatcctgcagtgtatagtatggtgcacatgaa 4375
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4192 TTCCCAGCAGCAGTTTGCCCACCAGGGGAATCCTGCAGTGTATAGTATGGTGCACATGAA 4251

Qy   4376 tggcagcagtggtcacatgggacagatgaacatgaaccccatgcccatgtctggcatgcc 4435
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4252 TGGCAGCAGTGGTCACATGGGACAGATGAACATGAACCCCATGCCCATGTCTGGCATGCC 4311

Qy   4436 tatgggtcctgatcagaaatactgctgacatctctgcaccaggacctcttaaggaaacca 4495
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4312 TATGGGTCCTGATCAGAAATACTGCTGACATCTCTGCACCAGGACCTCTTAAGGAAACCA 4371

Qy   4496 ctgtacaaatgacactgcactaggattattgggaaggaatcattgttccaggcatccatc 4555
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4372 CTGTACAAATGACACTGCACTAGGATTATTGGGAAGGAATCATTGTTCCAGGCATCCATC 4431

Qy   4556 ttggaagaaaggaccagctttgagctccatcaagggtattttaagtgatgtcatttgagc 4615
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4432 TTGGAAGAAAGGACCAGCTTTGAGCTCCATCAAGGGTATTTTAAGTGATGTCATTTGAGC 4491

Qy   4616 agga 4619
          ||||
Db   4492 AGGA 4495
```

```
RESULT   5
AF000581
LOCUS       AF000581     4609 bp    mRNA            ROD       06-MAR-1998
DEFINITION  Mus musculus p300/CBP/Co-Integrator protein mRNA, complete cds.
ACCESSION   AF000581
VERSION     AF000581.1  GI:2934905
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 4609)
  AUTHORS   Torchia,J., Rose,D.W., Inostroza,J., Kamei,Y., Westin,S.,
            Glass,C.K. and Rosenfeld,M.G.
  TITLE     The transcriptional co-activator p/CIP binds CBP and mediates
            nuclear-receptor function
  JOURNAL   Nature 387 (6634), 677-684 (1997)
  MEDLINE   97336097
REFERENCE   2  (bases 1 to 4609)
  AUTHORS   Torchia,J. and Rosenfeld,M.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-APR-1997) Medicine, Howard Hughes Medical Institute
            at the University of California, San Diego, 9500 Gilman Drive,
            C.M.M. 345, La Jolla, CA 92093-0648, USA
REFERENCE   3  (bases 1 to 4609)
  AUTHORS   Torchia,J. and Rosenfeld,M.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-MAR-1998) Medicine, Howard Hughes Medical Institute
            at the University of California, San Diego, 9500 Gilman Drive,
```

June 12, 1997

```
            C.M.M. 345, La Jolla, CA 92093-0648, USA
  REMARK    Sequence update by submitter
COMMENT     On Mar 6, 1998 this sequence version replaced gi:2213814.
FEATURES             Location/Qualifiers
     source          1. .4609
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
     CDS             110. .4306
                     /note="nuclear receptor coactivator protein; p/CIP"
                     /codon_start=1
                     /product="p300/CBP/Co-Integrator protein"
                     /protein_id="AAC05020.1"
                     /db_xref="GI:2934906"
                     /translation="MSGLGESSLDPLAAESRKRKLPCDAPGQGLVYSGEKWRREQESK
                     YIEELAELISANLSDIDNFNVKPDKCAILKETVRQIRQIKEQGKTISSDDDVQKADVS
                     STGQGVIDKDSLGPLLLQALDGFLFVVNRDGNIVFVSENVTQYLQYKQEDLVNTSVYS
                     ILHEPRRKDFLNTYQNPQLMEFLGLMRTRDKKAPYILIVRMLMKTHDILEDVNASPET
                     RQRYETMQCFALSQPRAMLEEGEDLQCCMICVARRVTAPFPSSPESFITRHDLSGKVV
                     NIDTNSLRSSMRPGFEDIIRRCIQRFFSLNDGQSWSQKRHYQEAYVHGHAETPVYRFS
                     LADGTIVSAQTKSKLFRNPVTNDRHGFISTHFLQREQNGYRPNPIPQDKGIRPPAAGC
                     GVSMSPNQNVQMMGSRTYGVPDPSNTGQMGGARYGASSSVASLTPGQSLQSPSSYQNS
                     SYGLSMSSPPHGSPGLGPNQQNIMISPRNRGSPKMASHQFSPAAGAHSPMGPSGNTGS
                     HSFSSSSLSALQAISEGVGTSLLSTLSSPGPKLDNSPNMNISQPSKVSGQDSKSPLGL
                     YCEQNPVESSVCQSNSRDPQVKKESKESSGEVSETPRGPLESKGHKKLLQLLTCSSDD
                     RGHSSLTNSPLDPNCKDSSVSVTSPSGVSSSTSGTVSSTSNVHGSLLQEKHRILHKLL
                     QNGNSPAEVAKITAEATGKDTSSTASCGEGTTRQEQLSPKKKENNALLRYLLDRDDPS
                     DVLAKELQPQADSGDSKLSQCSCSTNPSSGQEKDPKIKTETNDEVSGDLDNLDAILGD
                     LTSSDFYNNPTNGGHPGAKQQMFAGPSSLGLRSPQPVQSVRPPYNRAVSLDSPVSVGS
                     GPPVKNVSAFPGLPKQPILAGNPRMMDSQENYGANMGPNRNVPVNPTSSPGDWGLANS
                     RASRMEPLASSPLGRTGADYSATLPRPAMGGSVPTLPLRSNRLPGARPSLQQQQQQQQ
                     QQQQQQQQQQQQQQQMLQMRTGEIPMGMGVNPYSPAVQSNQPGSWPEGMLSMEQGPHG
                     SQNRPLLRNSLDDLLGPPSNAEGQSDERALLDQLHTFLSNTDATGLEEIDRALGIPEL
                     VNQGQALESKQDVFQGQEAAVMMDQKAALYGQTYPAQGPPLQGGFNLQGQSPSFNSMM
                     GQISQQGSFPLQGMHPRAGLVRPRTNTPKQLRMQLQQRLQGQQFLNQSRQALEMKMEN
                     PAGTAVMRPMMPQAFFNAQMAAQQKRELMSHHLQQQRMAMMMSQPQPQAFSPPPNVTA
                     SPSMDGVLAGSAMPQAPPQQFPYPANYGMGQPPEPAFGRGSSPPSAMMSSRMGPSQNA
                     MVQHPQPTPMYQPSDMKGWPSGNLARNGSFPQQQFAPQGNPAAYNMVHMNSSGGHLGQ
                     MAMTPMPMSGMPMGPDQKYC"
BASE COUNT     1220 a   1242 c   1191 g    955 t      1 others
ORIGIN
```

```
  Query Match           40.5%;  Score 2767.6;  DB 12;  Length 4609;
  Best Local Similarity 78.0%;  Pred. No. 0;
  Matches 3630;  Conservative    0;  Mismatches  820;  Indels  201;  Gaps   17;

Qy     92 ggcggcgggaggatcaaaa-tacttgctggatggtggactcagagaccaataaaaataaa 150
Db      1 GGCGGCGAACGGATCAAAAGAATTTGCTGAACAGTGGACTCCGAGATCGGTAAAACGAAC 60

Qy    151 ctgcttgaacatcctttgactggttagccagttgctgatgtatattcaagatgagtggat 210
Db     61 TCTTCCCTGCCCTTCCTGAACAGCT-GTCAGTTGCTGATCTGTGATCAGGATGAGTGGAC 119

Qy    211 taggagaaaac---ttggatccactggccagtgattcacgaaaacgcaaattgccatgtg 267
Db    120 TAGGCGAAAGCTCTTTGGATCCGCTGGCCGCTGAGTCTCGGAAACGCAAACTGCCCTGTG 179

Qy    268 atactccaggacaaggtcttacctgcagtggtgaaaaacggagacgggagcaggaaagta 327
Db    180 ATGCCCCAGGACAGGGGCTTGTCTACAGTGGTGAGAAGTGGCGACGGGAGCAGGAGAGCA 239

Qy    328 aatatattgaagaattggctgagctgatatctgccaatcttagtgatattgacaatttca 387
Db    240 AGTACATAGAGGAGCTGGCAGAGCTCATCTCTGCAAATCTCAGCGACATCGACAACTTCA 299

Qy    388 atgtcaaaccagataaatgtgcgattttaaaggaaacagtaagacagatacgtcaaataa 447
Db    300 ATGTCAAGCCAGATAAATGTGCCATCCTAAAGGAGACAGTGAGACAGATACGGCAAATAA 359

Qy    448 aagagcaaggaagaaactatttccaatgatgatgatgttcaaaaagccgatgtatcttcta 507
Db    360 AAGAACAAGGAAAAACTATTTCCAGTGATGATGATGTTCAAAAAGCTGATGTGTCTTCTA 419

Qy    508 cagggcagggagttattgataaagactccttaggaccgcttttacttcaggcattggatg 567
Db    420 CAGGGCAGGGAGTCATTGATAAAGACTCTTTAGGACCGCTTTTACTACAGGCACTGGATG 479

Qy    568 gtttcctatttgtggtgaatcgagacggaaacattgtatttgtatcagaaaatgtcacac 627
Db    480 GTTTCCTGTTTGTGGTGAATCGAGATGGAAACATTGTATTCGTGTCAGAAAATGTCACAC 539

Qy    628 aatacctgcaatataagcaagaggacctggttaacacaagtgtttacaatatcttacatg 687
Db    540 AGTATCTGCAGTACAAGCAGGAGGACCTGGTTAACACAAGTGTCTACAGCATCTTACATG 599

Qy    688 aagaagacagaaaggatttttcttaagaatttaccaaaatctacagttaatggagtttcct 747
Db    600 AGCCAAGACGGAAGGA-TTTCTTAAACACTTACCAAAATCCACAGTTAATGGAGTTTCTT 658

Qy    748 ggacaaatgagacccaaagacaaaaaagccatacatttaattgccgtatgt-tgatgaaa 806
Db    659 GGACTAATGAGAACCAGAGACAAAAAAGCCCCATACATTTAATTGTCCGTATGTTGATG 718

Qy    807 acaccacatgatattctggaagacataaacgccagtcctgaaatgcgccagagatatgaa 866
Db    719 AAAACACACGACATTTTGGAAGACGTGAATGCCAGTCCCGAAACGCGCCAGAGATATGAA 778

Qy    867 acaatgcagtgctttgccctgtctcagccacgagctatgatggaggaaggggaagatttg 926
Db    779 ACAATGCAGTGCTTTGCCCTGTCTCAGCCTCGCGCTATGCTGGAAGAAGGAGAAGACTTG 838

Qy    927 caatcttgtatgatctgtgtggcacgccgcattactacaggagaaagaacatttccatca 986
Db    839 CAGTGCTGTATGATCTGCGTGGCTCGCCGCGTGACTGC---------GCCATTCCCATCC 889

Qy    987 aaccctgagagctttattaccagacatgatctttcaggaaaggttgtcaatatagataca 1046
Db    890 AGTCCCGAGAGCTTTATTACCAGACATGACCTTTCCGGAAAGGTTGTCAATATAGATACA 949

Qy   1047 aattcactgagatcctccatgaggcctggctttgaagatataatccgaaggtgtattcag 1106
Db    950 AACTCACTTAGATCTTCCATGAGGCCTGGCTTTGAAGACATAATCCGAAGATGTATCCAG 1009

Qy   1107 agatttttttagtctaaatgatgggcagtcatggtcccagaaacgtcactatcaagaagct 1166
Db   1010 AGGTTCTTCAGTCTGAATGATGGGCAGTCATGGTCCCAGAAGCGTCACTATCAAGAAGCT 1069

Qy   1167 tatcttaatggccatgcagaaaccccagtatatcgattctcgttggctgatggaactata 1226
Db   1070 TATGTTCATGGCCACGCAGAGACCCCCGTGTATCGTTTCTCCTTGGCTGATGGAACTATT 1129

Qy   1227 gtgactgcacagacaaaaagcaaactcttccgaaatcctgtaacaaatgatcgacatggc 1286
Db   1130 GTGAGTGCGCAGACAAAAAGCAAACTCTTCCGCAATCCTGTAACGAATGATCGTCACGGC 1189

Qy   1287 tttgtctcaacccacttccttcagagagaacagaatggatatagaccaaacccaaatcct 1346
Db   1190 TTCATCTCGACCCACTTTCTTCAGAGAGAACAGAATGGATACAGACCAAACCCAATCCCG 1249

Qy   1347 gttggacaaggattagaccacctatggctggatgcaacagttcggtaggcggcatgagt 1406
Db   1250 CAGGACAAAGGCATCCGACCTCCTGCAGCAGGGT--------------GTGGCGTGAGC 1294

Qy   1407 atgtcgccaaaccaaggcttacagatgccgagcagcagggcctatggcttggcagaccct 1466
Db   1295 ATGTCTCCAAATCAGAATGTACAGATGATGGGCAGCCGGACCTATGGCGTGCCAGACCCC 1354

Qy   1467 agcaccacaggaggcagatgagtggagctaggtatgggggttccagtaacatagcttcattg 1526
Db   1355 AGCAACACAGGGCAGATGGGTGGAGCTAGGTACGGGGCTTCTAGTAGCGTAGCCTCACTG 1414

Qy   1527 acccctgggcaaggcatgcaatcaccatcttcctaccagaacaacaactatgggctcaac 1586
Db   1415 ACGCCAGGACAAAGCCTACAGTCGCCATCTTCCTATCAGAACAGCAGCTATGGGCTCAGC 1474
```

```
Qy   1587 atgagtagccccccacatgggagtcctggtcttgccccaaaccagcagaatatcatgatt 1646
Db   1475 ATGAGCAGTCCCCCCACGGCAGTCCTGGTCTTGGCCCCAACCAGCAGAACATCATGATT 1534

Qy   1647 tctcctcgtaatcgtgggagtccaaagatagcctcacatcagttttctcctgttgcaggt 1706
Db   1535 TCCCCTCGGAATCGTGGCAGCCCAAAGATGGCCTCCCACCAGTTCTCTCCTGCTGCAGGT 1594

Qy   1707 gtgcactctcccatggcatcttctggcaatactgggaaccacagcttttccagcagctct 1766
Db   1595 GCACACTCACCCATGGGACCTTCTGGCAACACAGGGAGCCACAGCTTTTCTAGCAGCTCC 1654

Qy   1767 ctcagtgccctgcaagccatcagtgaaggtgtggggacttccctttttatctactctgtca 1826
Db   1655 CTCAGTGCCTTGCAAGCCATCAGTGAAGGCGTGGGGACCTCTCTTTTATCTACTCTGTCC 1714

Qy   1827 tcaccaggccccaaattggataactctcccaatatgaatattacccaaccaagtaaagta 1886
Db   1715 TCACCAGGCCCCAAACTGGATAATTCTCCCAATATGAATATAAGCCAGCCAAGTAAAGTG 1774

Qy   1887 agcaatcaggattccaagagtcctctgggcttttattgcgaccaaaatccagtggagagt 1946
Db   1775 AGTGGTCAGGACTCTAAGAGCCCCCTAGGCTTATACTGTGAACAGAATCCAGTGGAGAGT 1834

Qy   1947 tcaatgtgtcagtcaaatagcagagatcacctcagtgacaaagaaagtaaggagagcagt 2006
Db   1835 TCAGTGTGTCAGTCAAACAGCAGAGATCCCCAAGTGAAAAAAGAAAGCAAGGAGAGCAGT 1894

Qy   2007 gttgaggggggcagagaatcaaaggggtcctttggaaagcaaaggtcataaaaaattactg 2066
Db   1895 GGGGAGGTGTCAGAGACGCCCAGGGGACCTCTGGAAAGCAAAGGCCACAAGAAACTGCTG 1954

Qy   2067 cagttacttacctgttcttctgatgaccggggtcattcctccttgaccaactccccccta 2126
Db   1955 CAGTTACTCACGTGCTCCTCCGACGACCGAGGCCATTCCTCCTTGACCAACTCTCCCCTG 2014

Qy   2127 gattcaagttgtaaagaatcttctgttagtgtcaccagcccctctggagtctcctcctct 2186
Db   2015 GATCCAAACTGCAAAGACTCTTCCGTTAGTGTCACCAGCCCCTCTGGAGTGTCCTCCTCA 2074

Qy   2187 acatctggaggagtatcctctacatccaatatgcatgggtcactgttacaagagaagcac 2246
Db   2075 ACATCAGGGACAGTGTCTTCCACCTCCAATGTGCATGGGTCTCTGTTGCAAGAGAAACAC 2134

Qy   2247 cggattttgcacaagttgctgcagaatgggaattcaccagctgaggtagccaagattact 2306
Db   2135 CGGATTTTGCACAAGTTGCTGCAGAATGGCAACTCCCCAGCGGAGGTCGCCAAGATCACT 2194

Qy   2307 gcagaagccactgggaaagacaccagcagtataacttcttgtggggacggaaatgttgtc 2366
Db   2195 GCAGAGGCCACTGGGAAGGACACGAGCAGCACTGCTTCCTGTGGAGAGGGGA---CAACC 2251

Qy   2367 aagcaggagcagctaagtcctaagaagaaggagaataatgcacttcttagatacctgctg 2426
Db   2252 AGGCAGGAGCAGCTGAGTCCTAAGAAGAAGGAGAATAATGCTCTGCTTAGATACCTGCTG 2311

Qy   2427 gacagggatgatcctagtgatgcactctctaaagaactacagccccaagtggaaggagtg 2486
Db   2312 GACAGGGATGACCCCAGTGATGTGCTTGCCAAAGAGCTGCAGCCCCAGGCCGACAGTGGG 2371

Qy   2487 gataataaaatgagtcagtgcaccagctccaccattcctagctcaagtcaagagaaagac 2546
Db   2372 GACAGTAAACTGAGTCAGTGCAGCTGCTCCACCAATCCCAGCTCTGGCCAAGAGAAAGAC 2431

Qy   2547 cctaaaattaagacagagacaagtgaagagggatctggagacttggataatctagatgct 2606
Db   2432 CCCAAAATTAAGACCGAGACGAACGACGAGGTATCGGGAGACCTGGATAATCTAGATGCC 2491

Qy   2607 attcttggtgatctgactagttctgacttttacaataattccatatcctcaaatggtagt 2666
Db   2492 ATTCTTGGAGATTTGACCAGTTCTGACTTCTACAACAATCCTA------CAAATGGCGGT 2545

Qy   2667 catctggggactaagcaacaggtgtttcaaggaactaattctctgggtttgaaaagttca 2726
Db   2546 CACCCAGGGGCCAAACAGCAGATGTTTGCAGGACCGAGTTCTCTGGGTTTGCGAAGTCCA 2605

Qy   2727 cagtctgtgcagtctattcgtcctccatataaccgagcagtgtctctggatagccctgtt 2786
Db   2606 CAGCCTGTGCAGTCTGTTCGTCCTCCATATAACCGAGCGGTGTCTCTGGATAGCCCTGTG 2665

Qy   2787 tctgttggctcaagtcctccagtaaaaaatatcagtgctttccccatgttaccaaagcaa 2846
Db   2666 TCTGTTGGCTCAGGTCCGCCAGTGAAGAATGTCAGTGCTTTCCCTGGGTTACCAAAACAG 2725

Qy   2847 cccatgttgggtgggaatccaagaatgatggatagtcaggaaaattatggctcaagtatg 2906
Db   2726 CCCATACTGGCTGGGAATCCAAGAATGATGGATAGTCAGGAGAATTACGGTGCCAACATG 2785

Qy   2907 ggtgggccaaaccgaaatgtgactgtgactcagactccttcctcaggagactggggctta 2966
Db   2786 ---GGCCCAAACAGAAATGTTCCTGTGAATCCGACTTCCTCCCCCGGAGACTGGGGCTTA 2842

Qy   2967 ccaaactcaaaggccggcagaatggaacctatgaattcaaactccatgggaagaccagga 3026
Db   2843 GCTAACTCAAGGGCCAGCAGAATGGAGCCTCTGGCATCAAGTCCCCTGGGAAGAACTGGA 2902

Qy   3027 ggagattataatacttctttacccagacctgcactgggtggctctattcccacattgcct 3086
Db   2903 GCCGATTACAGTGCCACTTTACCCAGACCTGCCATGGGGGGCTCTGTGCCTACCTTGCCA 2962

Qy   3087 cttcggtctaatagcataccaggtgcgagaccagtatt-------------------- 3124
Db   2963 CTTCGTTCTAATCGACTGCCAGGTGCAAGACCATCGTTGCAGCAACAGCAGCAGCAACAG 3022

Qy   3125 ---------------------------------gcaacagcagcagcagatgcttcaaatg 3152
Db   3023 CAGCAACAGCAACAACAACAGCAGCAACAGCAGCAACAGCAGCAGATGCTTCAAATG 3082

Qy   3153 aggcctggtgaaatccccatgggaatggggggctaatccctatggccaagcagcagcatct 3212
Db   3083 AGAACTGGTGAGATTCCCATGGGAATGGGAGTCAATCCNTATAGCCCAGCAGTGCAGTCT 3142

Qy   3213 aaccaactgggttcctggcccgatggcatgttgtccatggaacaagtttctcatggcact 3272
Db   3143 AACCAACCAGGTTCCTGGCCAGAGGGCATGCTCTCTATGGAACAAGGTCCTCACGGGTCT 3202

Qy   3273 caaaataggcctcttcttaggaattccctggatgatcttgttgggccaccttccaacctg 3332
Db   3203 CAAAATAGGCCTCTTCTTAGAAACTCTCTGGATGATCTGCTTGGGCCACCTTCTAACGCA 3262

Qy   3333 gaaggccagagtgacgaaagagcattattggaccagctgcacactcttctcagcaacaca 3392
Db   3263 GAGGGCCAGAGTGACGAGAGAGCTCTGCTGGACCAGCTGCACACATTCCTGAGCAACACA 3322

Qy   3393 gatgccacaggcctggaagaaattgacagagctttgggcattcctgaacttgtcaatcag 3452
Db   3323 GATGCCACAGGTCTGGAGGAGATCGACAGGGCCTTGGGAATTCCTGAGCTCGTGAATCAG 3382

Qy   3453 ggacaggcattagagcccaaacaggatgctttccaaggccaagaagcagcagtaatgatg 3512
Db   3383 GGACAAGCTTTGGAGTCCAAACAGGATGTTTTCCAAGGCCAAGAAGCAGCAGTAATGATG 3442

Qy   3513 gatcagaaggcaggattatatggacagacatacccagcacaggggcctccaatgcaagga 3572
Db   3443 GATCAGAAGGCTGCACTATATGGACAGACATACCCAGCTCAGGGTCCTCCCCTTCAAGGA 3502

Qy   3573 ggctttcatcttcagggacaatcaccatcttttaactctatgatgaatcagatgaaccag 3632
Db   3503 GGCTTTAACCTTCAGGGACAGTCACCATCGTTTAACTCTATGATGGGTCAGATTAGCCAG 3562

Qy   3633 caaggcaatttcctctccaaggaatgcacccacgagccaacatcatgagaccccggaca 3692
Db   3563 CAAGGCAGCTTCCTCTGCAAGGCATGCATCCTAGAGCCGGCCTCGTGAGACCAAGGACC 3622

Qy   3693 aacaccccccaagcaacttagaatgcagcttcagcagaggctgcagggccagcagtttttg 3752
```

```
Db   3623 AACACCCCGAAGCAGCTGAGAATGCAGCTTCAGCAGAGGCTACAGGGCCAGCAGTTTTTA 3682

Qy   3753 aatcagagccgacaggcacttgaattgaaaatggaaaaccctactgctggtggtgctgcg 3812
          |||||||||||| |||||||||||| |||||||||| ||||      ||| |||  ||||
Db   3683 AATCAGAGCCGGCAGGCACTTGAAATGAAAATGGAGAACC------CTGCTGGCACTGCT 3736

Qy   3813 gtgatgaggcctatgatgcagccccagcagggttttcttaatgctcaaatggtcgcccaa 3872
          |||||||||||| ||||||      |  |||| |||  ||||||| |||||||  |||||
Db   3737 GTGATGAGGCCCATGATG------CCCCAGGCTTTCTTTAATGCCCAAATGGCTGCCCAG 3790

Qy   3873 cgcagcagagagctgctaagtcatcacttccgacaacagagggtggctatgatgatgcag 3932
          |  |   |||||||| | || ||||||  | |   || |||||| |||| ||||||||||
Db   3791 CAGAAACGAGAGCTGATGAGCCATCACCTGCAGCAGCAGAGGATGGCGATGATGATGTCA 3850

Qy   3933 cagcagcagcagcaacagcagcagcagcagcagcagcagcaacagcaacagcaa 3992
          || |  ||||
Db   3851 CAACCACAGC---------------------------------------------- 3860

Qy   3993 cagcaacagcagcaacagcagcaaacccaggccttcagcccacctcctaatgtgactgct 4052
                                    | |||||||||||||||||||| || || || ||
Db   3861 ------------------------CTCAGGCCTTCAGCCCACCTCCCAACGTCACCGCC 3895

Qy   4053 tcccccagcatggatgggcttttggcaggacccacaatgccacaagctcctccgcaacag 4112
          |||||||||||||| ||| |||||||||||  |  ||||||| ||||| ||||| ||||||
Db   3896 TCCCCCAGCATGGACGGGGTTTTGGCAGGTTCAGCAATGCCGCAAGCCCCTCCACAACAG 3955

Qy   4113 tttccatatcaaccaaattatggaatgggacaacaaccagatccagcctttggtcgagtg 4172
          ||||||||||  | |||||||  |||||||||||||  ||||||  ||||||||||||||||
Db   3956 TTTCCATATCCAGCAAATTACGGAATGGGACAACCACCAGAGCCAGCCTTTGGTCGAGGC 4015

Qy   4173 tctagtcctcccaatgcaatgatgtcgtcaagaatgggtccctcccagaatcccatgatg 4232
          || ||||||||||  |||||||||||  ||||||||||  || |||||||||  ||||| ||
Db   4016 TCGAGTCCTCCCAGTGCAATGATGTCATCAAGAATGGGGCCTTCCCAGAATGCCATGGTG 4075

Qy   4233 caacacccgcaggctgcatccatctatcagtcctcagaaatgaagggctggccatcagga 4292
          || || ||  |||  |  || |||| |||||| | |||||  |||||||| |||||  |||||
Db   4076 CAGCATCCTCAGCCCACACCCATGTATCAGCCTTCAGATATGAAGGGGTGGCCGTCAGGG 4135

Qy   4293 aatttggccaggaacagctccttttcccagcagcagtttgcccaccaggggaatcctgca 4352
          ||  ||||||||||  |||||||  ||||||||||||||||  | |||||||||  ||||||
Db   4136 AACCTGGCCAGGAATGGCTCCTTCCCCCAGCAGCAGTTTGCTCCCCAGGGGAACCCTGCA 4195

Qy   4353 gtgtatagtatggtgcacatgaatggcagcagtggtcacatgggacagatgaacatgaac 4412
          |  || |  |||||||| |||||  |||||  |||| ||| ||||||||||  ||||| |
Db   4196 GCCTACAACATGGTGCATATGAACAGCAGCGGTGGGCACTTGGGACAGATGGCCATGACC 4255

Qy   4413 cccatgcccatgtctggcatgcctatgggtcctgatcagaaatactgctgacatctctgc 4472
          ||||||||||||||||||||||| ||||| || ||||||||||||||||||||||||||
Db   4256 CCCATGCCCATGTCTGGCATGCCCATGGGCCCCGATCAGAAATACTGCTGACATCTCCCT 4315

Qy   4473 accaggacctcttaaggaaaccactgtacaaatgacactgcactaggattattgggaagg 4532
          |  ||||        ||    || |  | |  ||| | |  |  |  |  ||  |
Db   4316 AGTGGGAC-------TGACTGTACAGATGACACTGCACAGGATCATCAGGACGTGGCGGC 4368

Qy   4533 aatcattgttccaggcatccatcttggaagaaaggaccagctttgagctccatcaagggt 4592
            |   | ||  | |||||||  ||||||||  |||| |         |||   | ||   |
Db   4369 GAGTCATTGTCTAAGCATCCAGCTTGGAAGCAAGGCC--------AGCGTGACCAGCAGC 4420

Qy   4593 attttaagtgatgtcatttgagcaggactggattttaagccgaagggcaatatctacgtg 4652
               |  ||| ||||||||||||||  |||| | |   || |||| ||| | ||||| ||
Db   4421 GGGGTCTGTGCTGTCATTTGAGCAGAGCTGGGTCT--CGCTGAAGCGCACTGTCTACCTG 4478

Qy   4653 ttttcccccctccttctgctgtgtatcatg 4683
             |  | ||    |  ||  || |  ||
Db   4479 ATGCCCTGCCTCTGTGTGGCAAGGTGTTCTG 4509


RESULT   6
HSU80737
LOCUS       HSU80737     3487 bp    mRNA            PRI       18-DEC-1997
DEFINITION  Homo sapiens CAGH16 mRNA, complete cds.
ACCESSION   U80737
VERSION     U80737.1  GI:2565049
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3487)
  AUTHORS   Margolis,R.L., Abraham,M.R., Gatchell,S.B., Li,S.H., Kidwai,A.S.,
            Breschel,T.S., Stine,O.C., Callahan,C., McInnis,M.G. and Ross,C.A.
  TITLE     cDNAs with long CAG trinucleotide repeats from human brain
  JOURNAL   Hum. Genet. 100 (1), 114-122 (1997)
  MEDLINE   97369492
REFERENCE   2  (bases 1 to 3487)
  AUTHORS   Margolis,R.L., Abraham,M.R., Gatchell,S.B., Li,S.H., Kidwai,A.S.,
            Breschel,T.S., Stine,O.C., Callahan,C., McInnis,M.G. and Ross,C.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-DEC-1996) Psychiatry, Johns Hopkins Univ. Sch. of
            Med., 600 N. Wolfe Street, Meyer 2-181, Baltimore, MD 21205, USA
FEATURES             Location/Qualifiers
     source          1. .3487
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /map="20q13.13; between D20S891 and D20S109"
                     /dev_stage="fetus"
                     /tissue_type="brain"
                     /note="single strand sequence"
     gene            1. .3487
                     /gene="CAGH16"
     CDS             710. .1690
                     /gene="CAGH16"
                     /note="similar to Hin-2 encoded by GenBank Accession
                     Number U19179; polyglutamine rich"
                     /codon_start=1
                     /product="CAGH16"
                     /protein_id="AAB91436.1"
                     /db_xref="GI:2565050"
                     /translation="MMDQKAGLYGQTYPAQGPPMQGGFHLQGQSPSFNSMMNQMNQQG
                     NFPLQGMHPRANIMRPRTNTPKQLRMQLQQRLQGQQFLNQSRQALELKMENPTAGGAA
                     VMRPMMQPQVSSQEGFLNAQMGAQRRRELLSHHFRQQRVAMMMQQQQQQQQQQQQQQQ
                     QQQQQQQQQQQQQEQPQVFSPPPNVTAPPAWMAXGRTHNATSSSATSFQYPTPNYGNG
                     DQPTKIQPLGRVSKSSQCNDVVKNGPSQNPMMQHPQAASIYQSSEMKGWPSGNLARNS
                     SFSQQQFAHQGNPAVYSMVHMNGSSGHMGQMNMNPMPMSGMPMGPDQKYC"
     misc_feature    1145. .1234
                     /gene="CAGH16"
                     /note="CAG/CAA region; glutamine rich region"
     repeat_region   1166. .1201
                     /rpt_type=tandem
                     /rpt_unit=CAG
BASE COUNT      924 a    781 c    775 g   1002 t      5 others
ORIGIN


  Query Match            36.4%;  Score 2485.4;  DB 66;  Length 3487;
  Best Local Similarity  96.6%;  Pred. No. 0;
  Matches 2758;  Conservative    2;  Mismatches   51;  Indels   45;  Gaps   20;

Qy   3426 ttgggcattcctgaacttgtcaatcagggacaggcattagagcccaaacaggatgctttc 3485
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    629 TAGGGCATTCCTGAACTTGTCAATCAGGGACAGGCATTAGAGCCCAAACAGGATGCTTTC 688

Qy   3486 caaggccaagaagcagcagtaatgatggatcagaaggcaggattatatggacagacatac 3545
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    689 CAAGGCCAAGAAGCAGCAGTAATGATGGATCAGAAGGCAGGATTATATGGACAGACATAC 748

Qy   3546 ccagcacaggggcctccaatgcaaggaggctttcatcttcagggacaatcaccatctttt 3605
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    749 CCAGCACAGGGGCCTCCAATGCAAGGAGGCTTTCATCTTCAGGGACAATCACCATCTTTT 808

Qy   3606 aactctatgatgaatcagatgaaccagcaaggcaattttcctctccaaggaatgcaccca 3665
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    809 AACTCTATGATGAATCAGATGAACCAGCAAGGCAATTTTCCTCTCCAAGGAATGCACCCA 868

Qy   3666 cgagccaacatcatgagaccccggacaaacacccccaagcaacttagaatgcagcttcag 3725
```